Measuring Hospital-Wide Mortality—Pitfalls and Potential

Simon J. Mackenzie, Don A. Goldmann, Rocco J. Perla, and Gareth J. Parry

Abstract: Risk-adjusted hospital-wide mortality has been proposed as a key indicator of system-level quality. Several risk-adjusted measures are available, and one—the hospital standardized mortality ratio (HSMR)—is publicly reported in a number of countries, but not in the United States. This paper reviews potential uses of such measures. We conclude that available methods are not suitable for interhospital comparisons or rankings and should not be used for pay-for-performance or value-based purchasing/payment. Hospital-wide mortality is a relatively imprecise, crude measure of quality, but disaggregation into condition- and service-line–specific mortality can facilitate targeted improvement efforts. If tracked over time, both observed and expected mortality rates should be monitored to ensure that apparent improvement is not due to increasing expected mortality, which could reflect changes in case mix or coding. Risk-adjusted mortality can be used as an initial signal that a hospital's mortality rate is significantly higher than statistically expected, prompting further inquiry.

Keywords
mortality performance improvement/quality improvement
public reporting/transparency
quality measures

The Francis Report (Francis, 2013) documented numerous failures of care and potentially avoidable deaths at the Mid Staffordshire NHS Trust in England. Prior to the report, there was data, readily available, indicating higher than expected mortality and this has renewed debate about the use of risk-adjusted mortality rates as an indicator of system-level quality (Mid Staffordshire NHS Foundation Trust Inquiry, 2010). Hospital-wide mortality measures are appealing because death is an unambiguous event and reporting is usually mandated by law. Risk-adjusted hospital-wide mortality has been proposed to enable hospitals to gauge the success of improvement initiatives over time despite changes in case mix, facilitate interhospital comparisons and inform consumers, and help regulatory or governmental agencies detect hospitals with exceptionally high (or low) mortality rates. Hospital-wide mortality has also been considered as a measure for pay-for-performance and value-based purchasing/payment, joining a growing array of measures (such as readmissions and hospital-acquired infections) that have been deployed in the United States. A risk-adjusted hospital-wide mortality measure, the hospital standardized mortality ratio (HSMR), has been available publicly for more than a decade in England and in several other countries, but not in the United States (Jarman et al., 2010). In the U.S., hospitals have access to HSMR, and several other risk-adjusted hospital-wide mortality measures for their own use. In this paper, we describe some of the pitfalls and potential utility of developing and applying hospital-wide mortality measures.

Method

Drawing on key publications, reports in the field, and published presentations from an expert meeting sponsored by the Agency for Healthcare Research and Quality and the Institute for Healthcare Improvement, we focus on model development, comparing performance, identifying outliers, and catalyzing improvement (Agency for Healthcare Research and Quality, 2009). We also undertook an additional simple literature review to include papers published since that meeting. We searched PubMed from 2009 to 2014, using the following search terms: "Quality of care" and ("SHMI" or "HSMR" or "Hospital Standardized Mortality Ratio" or "Hospital-wide mortality"). Excluding 26 publications that were letters or correspondence, focused on condition-specific applications, or were beyond the technical scope of this paper, 17 papers published between 2009 and 2014 were identified. This paper presents a narrative summary of the pitfalls and potential of hospital-wide mortality measures, based on these sources.

Model Development

Within the United States, all available risk-adjusted hospital-wide mortality measures rely on administrative data for risk-adjustment. It is generally acknowledged that using only administrative data in risk-adjustment models is problematic, and including laboratory and clinical information in these models—as will be possible when databases incorporating electronic medical record data become more widespread—will improve their performance (Escobar et al., 2008). There is growing consensus that mortality rates that extend at least 30

Journal for Healthcare Quality
Vol. 00, No. 0, pp. 1–7

days beyond hospital discharge are preferable to those that include in-hospital mortality only, and that diagnosis- and service-line–specific mortality rates provide more focused, actionable data (Downar, Sibbald, & Lazar, 2010; Holloway & Quill, 2007).

Any performance measure should be developed and evaluated according to its intended use. A measure designed for a quality improvement department to assess the impact of its program over time might have different properties than one intended for a national pay-for-performance program. Fastidious risk adjustment is important if a measure is intended for comparisons, judgment, and payment, but probably is not necessary to track improvement within a hospital, especially if its case mix remains relatively stable over time. Nonetheless, a "one-size-fits-all" approach has been the norm, leading to measures that are appropriate to inform improvement being used, inappropriately, for judgment (Solberg, Mosser, & McDonald, 1997).

Comparing Performance: Measuring Mortality for Public Reporting, Interhospital Comparisons, and Payment

Not surprisingly, the use of hospital-wide mortality measures to compare, judge, and reimburse hospitals is particularly contentious (Scott, Brand, Phelps, Barker, & Cameron, 2011; van den Bosch, Kelder, & Wagner, 2011). Problems with previous attempts to publicly report hospital mortality have sharpened the debate about existing risk-adjustment methods and the degree to which hospital-wide mortality measures accurately reflect patient safety and quality and can be used to compare hospital performance (Black, 2010; van Gestel et al., 2012).

The "Standards for statistical models used for reporting of health outcomes" recommend seven attributes any risk-adjustment model should possess for use as a measure for judgment (Krumholz et al., 2006): definition of the patient sample, clinical coherence, data quality, designation of a reference time, standardized assessment of meaningful outcomes, appropriate analytical approach, and public disclosure. No published studies have evaluated hospital-wide mortality measures against these criteria. However, Shahian, Wolf, Iezzoni, Kirle, and Normand (2010) found that four hospital-wide mortality risk-adjustment methods had dramatically different performance characteristics when applied to a hospital data set in Massachusetts. They noted potential problems with all of the models; they found differences in the inclusion criteria for patients included in the models, and that the same hospital could be found to have higher-than-expected mortality by one method, but lower-than-expected by another. This demonstrates the hazards of using hospital-wide mortality measures to compare or rank hospitals.

More fundamentally, Shahian, Iezzoni, Meyer, Kirle, and Normand (2012) questioned whether hospital-wide mortality is a reliable indicator of quality of care and suggested focusing on a limited subgroup of diagnoses for which the link between mortality and quality is more plausible. Jha, Zhonghe, Orav, and Epstein (2005) cautioned that use of hospital-wide mortality as an indicator of overall quality might obscure important variations in mortality rates within the hospital, with high mortality and low mortality areas canceling each other out, yielding an overall average rate. Other investigators have suggested that the "signal-to-noise" ratio is too low for mortality to be a reliable measure of quality. Two studies based on medical record review suggested that only 7% of in-hospital deaths are potentially avoidable. Moreover, these studies conclude that these preventable deaths predominantly occur in patients who are admitted and treated, but had a poor prognosis and were expected to die soon (Hogan et al., 2012). In many cases, the authors concluded it was the timing of the death that was altered rather than its occurrence. Although these studies did not examine the potential contribution of systems improvement initiatives (e.g., communication and coordination, prompt rescue of deteriorating patients, strong safety culture) to mortality reduction, they suggest that it will be challenging to design programs likely to have a meaningful and measurable impact on avoidable hospital-wide mortality.

Even if a new risk-adjusted mortality measure was robust enough to satisfy critics, it still would have limited value for interhospital comparisons and rankings, or for value-based purchasing/payment that depends on comparative performance. Differences between hospitals' risk-adjusted mortality rates cannot be measured with sufficient precision to develop fine rankings of individual performance, potentially leading to misclassification. Research

shows that the great majority of hospitals in a typical ranking scheme will be statistically indistinguishable from one another (Parry, Gould, McCabe, & Tarnow-Mordi, 1998). Reporting the ranking of hospitals frequently results in defensive, contentious debates that distract from efforts to focus on improvements to reduce mortality (Lilford & Pronovost, 2010).

Risk-adjusted hospital-wide mortality measures do not provide for precise rankings of hospitals and have an unclear association with hospital-wide quality. Consequently, there is no clear basis for using such measures to improve hospital-wide quality in pay-for-performance or value-based purchasing.

Identifying Outliers: Using Risk-Adjusted Hospital-Wide Mortality to Detect Hospitals with Exceptionally High Mortality Rates

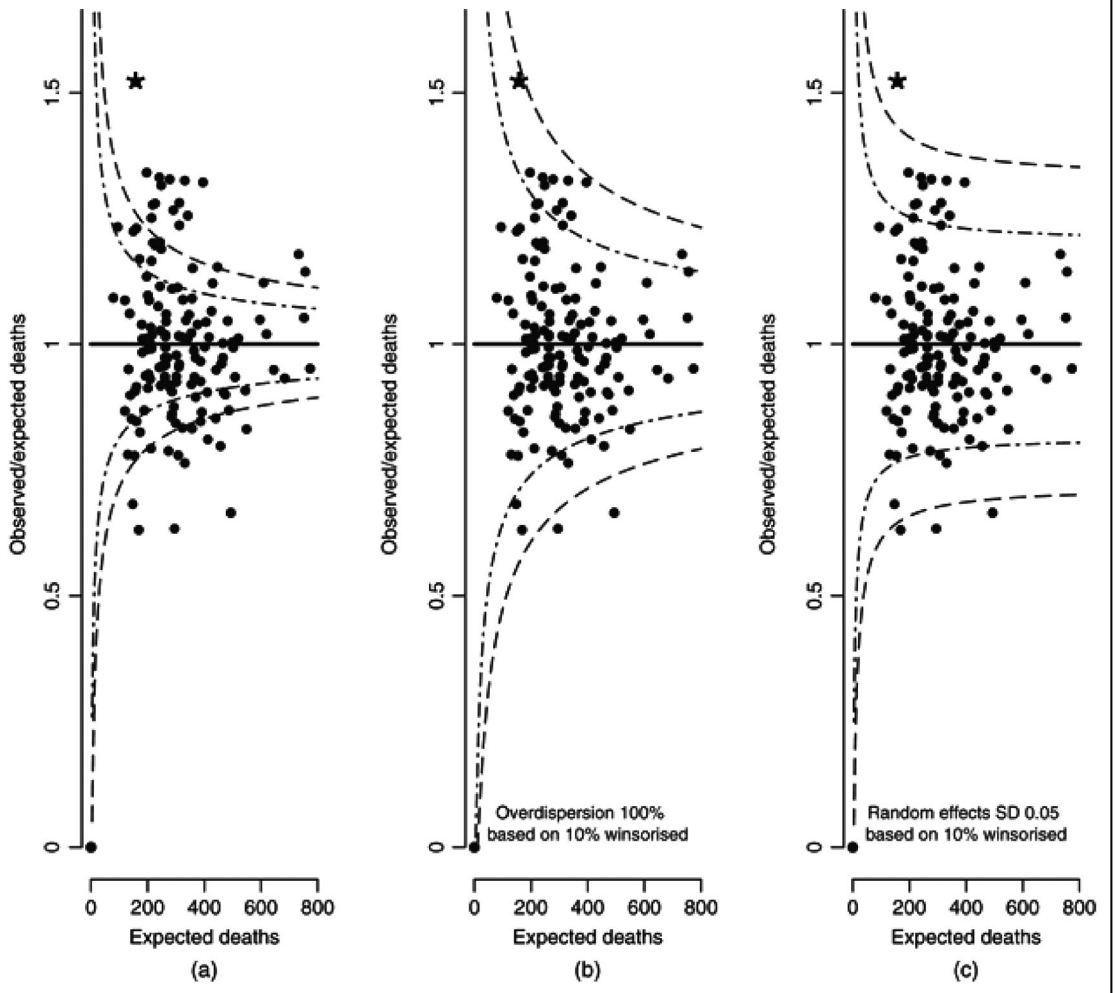
Although risk-adjusted hospital-wide mortality measures should not be used to compare hospitals or for value-based purchasing/payment, they could be useful as a signal to detect hospitals with statistically significant high risk-adjusted mortality rates (statistical outliers) that are unlikely to be due to subtle or unmeasured differences in case mix, especially if sustained over time or confirmed by another mortality measurement method. Funnel plots are a useful tool for identifying hospitals that are statistically indistinguishable from those that are statistical outliers. Funnel plots display mortality rates on the *y*-axis against the denominator (usually the number of admissions) on the *x*-axis (Campbell, Jacques, Fotheringham, Maheswaran, & Nicholl, 2012; Schulman, Spiegelhalter, & Parry, 2008). The plot includes limits, specified according to the desired purpose, that get narrower as the sample size increases. A value that lies outside these preset limits can be considered "above expectation," "significantly above expectation," or "substantially different from expectation" (Figure 1). Importantly, this is only a statistical test. It does not mean that hospitals within the limits provide good care, hospitals with low rates are outstanding, or hospitals with high rates are poor, it is simply a starting point for further inquiry. Other techniques, such as the cumulative sum (CUSUM) control chart, are useful in prompt detection of statistically significant changes in mortality (Grigg, Farewell, & Spiegelhalter, 2003; Jarman, Bottle, Aylin, & Brown, 2005).

The Francis report focused attention on the importance of having a public safety system in place to react promptly to warning signals that an institution may be providing poor-quality care, and offered a template for how to proceed when a statistical alarm is triggered. Building off the statistical approach described above, the risk-adjusted mortality rate can be viewed as a screening test to guide actions following the identification of an organization with a mortality rate that is substantially higher than expected (Healthcare Commission, 2009; Spiegelhalter et al., 2012). Following a signal, additional information would be requested confidentially from hospital leadership to explain the mortality data. Only following review of this information would a more in-depth external investigation be triggered. Such an evaluation would include site visits; interviews with leadership, staff, and patients; review of quality and safety data; and an independent medical record review to determine whether or not outlier hospitals do have system-wide quality issues. This approach has been adopted in both England (The Keogh Mortality Review, 2013) and Scotland (Healthcare Improvement Scotland, 2013). This process is analogous to a screening test; even when the test has high sensitivity (as required of a screening test) and relatively high specificity, careful follow-up of positive tests is required to differentiate true- from false-positives. Like any screening tool, such an approach will require testing and tracking of unintended consequences. Statistical thresholds based on sound epidemiological principles and thoughtful, disciplined review systems are needed to achieve a balance between protecting the public from organizations that are genuinely poor and protecting the reputation of hospitals that are not (Freemantle et al., 2013; Mohammed, Lilford, Rudge, Holder, & Stevens, 2013).

Improvement: Measuring Mortality to Catalyze and Monitor Improvement

Hospitals that initiate programs to reduce mortality will need a reliable measure to determine if improvement is occurring. Rigorous risk-adjustment is not necessary for this purpose if case mix has varied marginally over time (Richardson, Tarnow-Mordi, & Lee, 1999). Unadjusted mortality is unaffected by coding issues, is more timely than a risk-adjusted

Figure 1. Funnel Plots with (a) No, (b) Multiplicative, and (c) Additive Overdispersion, Showing the Observed and Expected Deaths from all Adult Emergency Admissions in the Third Quarter of 2006



Note. *, Mid Staffordshire Trust. This figure illustrates a variety of approaches for displaying data in the form of a funnel plot. The purpose of detecting outliers and what actions will occur following detection should be specified in advance. Reproduced with permission from: Spiegelhalterand colleagues (2012).

measure, and can be parsed into condition- and service-line–specific rates to inform targeted improvement efforts. Deaths, especially in patients who have a low risk of dying, should be studied carefully to further illuminate opportunities for improvement (Move Your Dot™: Measuring, Evaluating, and Reducing Hospital Mortality Rates [Part 1], 2003). For example, such reviews might show that patients who are at the end of life are dying in the intensive care unit rather than in a more appropriate setting. They might indicate that patients presenting with sepsis account for a disproportionate number of deaths, suggesting a need for more timely identification and initiation of evidence-based care for sepsis in the emergency department.

If a risk-adjusted measure is used to track improvement, it should be graphed along with its components—observed and expected mortality. The expected mortality is affected not just by changes in case mix, but also by coding (Jacques, Fotheringham, Campbell, & Nicholl, 2013; Song et al., 2010; Welch, Sharp, Gottlieb, Skinner, & Wennberg, 2011; Wennberg et al., 2013). Changes in coding over time alter expected mortality and produce a reduction

in adjusted mortality that may not reflect improvement in care. For example, in the United Kingdom, in "The North West Reducing Mortality Collaborative" run by the Advancing Quality Alliance, the reduction in risk-adjusted hospital-wide mortality was reported to be due to an increase in expected mortality following changes to hospital coding practices (Advancing Quality Alliance, 2011). To learn whether improvement is really occurring, hospitals should track unadjusted mortality and monitor the impact of changes in coding versus case mix on their adjusted mortality measure (Hawkes, 2010; Klugman, Allen, Benjamin, Fitzgerald, & Ettinger, 2010).

Whether or not risk-adjusted measures are used to monitor mortality trends, mortality should be viewed as just one component of a suite of measures designed to monitor a range of quality issues—for example, rates of hospital-acquired conditions; management and outcomes of common conditions, such as heart failure, stroke, and acute myocardial infarction; patient perceptions of care; efficiency; equity; and workforce safety.

Given the complexities of using risk-adjusted mortality measures to gauge the success of improvement efforts, hospitals should be cautious about claiming they have "saved" a specific number of lives. Even if it can be shown that changes in expected mortality are not attributable to changes in coding, it is important to determine whether an individual who did not die in the hospital subsequently died within a relatively short period of time following discharge, in which case a "life saved" would be a "death deferred." Most risk-adjusted mortality measures extend to 30 days postdischarge, but consideration should be given to examining mortality over a longer period of time—say, 90–180 days postdischarge.

Discussion and Conclusion

This paper summarizes the issues raised in reports, publications, and expert meeting summaries of hospital-wide mortality measurement, and offers recommendations on where there is potential to move the field forward. Given this is not a systematic review, there is potential for bias due to related studies not being included. Despite this limitation, we conclude that hospital-wide mortality is an imprecise measure of hospital quality, and should be supplemented by analysis of condition- or service-line–specific mortality and an examination of the causes of in-hospital death, which can facilitate targeted improvement efforts. Available data suggest that current hospital-wide mortality models available in the United States are unsuitable for making accurate comparisons of performance and should not be used to penalize hospitals or guide value-based models of purchasing/payment. We recommend that value-based models continue to focus on condition-specific measures. If risk-adjusted mortality is used to track improvement over time, its components—observed and expected mortality—should be tracked as well, with particular attention to whether or not change is due to changes in case mix and coding. In order to capture deaths that are attributable to hospital care but occur postdischarge, it is preferable to use a measure that extends at least 30 days beyond hospital discharge, as is the case for publicly reported mortality for acute myocardial infarction, heart failure, and pneumonia in the United States.

We recommend the use of available risk-adjusted hospital-wide mortality measures as a vehicle for further inquiry to detect and alert hospitals with exceptionally high mortality rates, especially if elevated rates persist over time. Recognizing that this alert is only a statistical signal to inform the need for deeper inquiry, we recommend creating an oversight system to provide confidential feedback to outlier hospitals. This feedback will inform expedited and timely internal reviews to evaluate the accuracy of the signal prior to public reporting and determine the need for external inquiry.

Currently, despite the availability and wide use of a number of risk-adjusted hospital-wide mortality measures, the United States does not have a system to rapidly and reliably detect possible high-mortality hospitals and to initiate an appropriate, timely inquiry. Similarly, such a system does not exist for publicly reported condition-specific mortality measures that include deaths within 30 days following discharge. The Francis report is a reminder that the United States needs to be more aggressive in using a wide variety of data—imperfections notwithstanding—as part of a deliberate, balanced, fair system to protect public safety.

Acknowledgments

The authors thank Matt Stiefel, Carolyn Clancy, and Jane Roessner for their valuable and critical feedback of previous versions of this manuscript.

References

Advancing Quality Alliance. (2011). North West Mortality Reduction Collaborative. Phase 1 Summary Report 2011. Retrieved February 25, 2014, from http://217.33.237.40/document_uploads/Rachel-documents/AQuA%20NW%20Reducing%20Mortality%20Report%20Phase%201%20output%20report%20June%202011_9b964.pdf.

Agency for Healthcare Research and Quality. (2009). Presentations from a November 2008 meeting to discuss issues related to mortality measures. Retrieved February 25, 2014, from www.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/mortality/.

Black, N. (2010). Assessing the quality of hospitals. *British Medical Journal*, *340*, 933–934. c2066.

Campbell, M. J., Jacques, R. M., Fotheringham, J., Maheswaran, R., & Nicholl, J. (2012). Developing a summary hospital mortality index: Retrospective analysis in English hospitals over five years. *BMJ*, *1*, 1–11. e1001. doi:10.1136/bmj.e1001

Downar, J., Sibbald, R., & Lazar, N. M. (2010). Ethical considerations for classifying patients as "palliative" when calculating hospital standardised mortality ratios. *Journal of Medical Ethics*, *36*, 387–390. doi:10.1136/jme.2010.035451

Escobar, G. J., Greene, J. D., Scheirer, P., Gardner, M., Draper, D., & Kipnis, P. (2008). Risk adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. *Medical Care*, *46*, 232–239.

Francis, R. (2013). Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London: House of Commons. Retrieved May 16, 2014, from www.midstaffspublicinquiry.com/report.

Freemantle, N., Richardson, M., Wood, J., Ray, D., Khosla, S., Sun,P., et al. (2013). Can we update the Summary Hospital Mortality Index (SHMI) to make a useful measure of the quality of hospital care? An observational study. *BMJ Open*, *3*, 1–8. e002018.

Grigg, O. A., Farewell, V. T., & Spiegelhalter, D. J. (2003). Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Statistical Methods in Medical Research*, *12*, 147–170.

Hawkes, N. (2010). Patient coding and the ratings game. *British Medical Journal*, *340*, c2153.

Healthcare Commission. (2009). Following up mortality "outliers": A review of the programme for taking action where data suggest there may be serious concerns about the safety of patients. London: Healthcare Commission. Retrieved February 25, 2014, from www.cqc.org.uk/_db/_documents/Following_up_mortality_outliers_200906054425.pdf.

Healthcare Improvement Scotland. (2013). Rapid review of the safety and quality of care for acute adult patients in NHS Lanarkshire. Retrieved February 25, 2014, from www.healthcareimprovementscotland.org/our_work/governance_and_assurance/programme_resources/nhs_lanarkshire__review.aspx.

Hogan, H., Healey, F., Neale, G., Thomson, R., Vincent, C., & Black, N. (2012). Preventable deaths due to problems in care in English acute hospitals: A retrospective case record review study. *BMJ Quality and Safety*, *21*, 737–745.

Holloway, R. G., & Quill, T. E. (2007). Mortality as a measure of quality: Implications for palliative and end-of-life care. *Journal of the American Medical Association*, *298*, 802–804.

Jacques, R. M., Fotheringham, J., Campbell, M. J., & Nicholl, J. (2013). Did hospital mortality in England change from 2005 to 2010? A retrospective cohort analysis. *BMC Health Services Research*, *13*, 216.

Jarman, B., Bottle, A., Aylin, P., & Brown, M. (2005). Monitoring changes in hospital standardised mortality ratios. *British Medical Journal*, *330*, 329.

Jarman, B., Pieter, D., van der Veen, A. A., Kool, R. B., Aylin, P., Bottle, A., et al. (2010). The hospital standardised mortality ratio: A powerful tool for Dutch hospitals to assess their quality of care? *Quality and Safety in Health Care*, *19*, 9–13. doi:10.1136/qshc.2009.032953

Jha, A. K., Zhonghe, L., Orav, E. J., & Epstein, A. M. (2005). Care in US Hospitals—The hospital quality alliance program. *New England Journal of Medicine*, *353*, 265–274.

Keogh, B. (2013). Review into the quality of care and treatment provided by 14 hospital trusts in England: Overview report. England: National Health Service. Retrieved February 25, 2014, from www.nhs.uk/NHSEngland/bruce-keogh-review/Documents/outcomes/keogh-review-final-report.pdf.

Klugman, R., Allen, L., Benjamin, E. M., Fitzgerald, J., & Ettinger, W., Jr. (2010). Mortality rates as a measure of quality and safety, "caveat emptor." *American Journal of Medical Quality*, *25*, 197–201.

Krumholz, H. M., Brindis, R. G., Brush, J. E., Cohen, D. J., Epstein, A. J., Furie, K., et al. (2006). Standards for statistical models used for reporting of health outcomes: An American Heart Association scientific statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group. *Circulation*, *113*, 456–462.

Lilford, R., & Pronovost, P. (2010). Using hospital mortality rates to judge hospital performance: A bad idea that just won't go away. *British Medical Journal*, *340*, 955–957. c2016.

Mid Staffordshire NHS Foundation Trust Inquiry. (2010). Independent inquiry into care provided by Mid Staffordshire NHS Foundation Trust, January 2005—March 2009. Retrieved February 25, 2014, from www.midstaffsinquiry.com/assets/docs/Inquiry_Report-Vol1.pdf.

Mohammed, M. A., Lilford, R., Rudge, G., Holder, R., & Stevens, A. (2013). The findings of the Mid Staffordshire Inquiry do not uphold the use of hospital standardized mortality ratios as a screening test for "bad" hospitals. *Quarterly Journal of Medicine*, *106*, 849–854. doi:10.1093/qjmed/hct101

Move Your Dot™: Measuring, Evaluating, and Reducing Hospital Mortality Rates (Part 1). (2003). *IHI Innovation Series white paper*. Boston, MA: Institute for Healthcare Improvement. Retrieved February 25, 2014, from www.ihi.org/resources/Pages/IHIWhitePapers/MoveYourDotMeasuringEvaluatingandReducingHospitalMortalityRates.aspx.

Parry, G. J., Gould, C. R., McCabe, C. J., & Tarnow-Mordi, W. O. (1998). Annual league tables of hospital mortality in neonatal intensive care: A longitudinal study. *British Medical Journal*, *316*, 1931–1935.

Richardson, D., Tarnow-Mordi, W. O., & Lee, S. K. (1999). Risk-adjustment for quality improvement. *Pediatrics*, *103*(Suppl), 255–265.

Schulman, J., Spiegelhalter, D. J., & Parry, G. (2008). How to interpret your dot: Decoding the message of clinical performance indicators. *Journal of Perinatology*, *28*, 588–596.

Scott, I. A., Brand, C. A., Phelps, G. E., Barker, A. L., & Cameron, P. A. (2011). Using hospital standardised

mortality ratios to assess quality of care–proceed with extreme caution. *Medical Journal of Australia, 194*, 645–648.
Shahian, D. M., Iezzoni, L. I., Meyer, G. S., Kirle, L., & Normand, S. L. (2012). Hospital-wide mortality as a quality metric: Conceptual and methodological challenges. *American Journal of Medical Quality, 27*, 112–123.
Shahian, D. M., Wolf, R. E., Iezzoni, L., Kirle, L., & Normand, S. L. (2010). Variability in the measurement of hospital-wide mortality rates. *New England Journal of Medicine, 363*, 2530–2539.
Solberg, L., Mosser, G., & McDonald, S. (1997). The three faces of performance improvement: Improvement, accountability and research. *Joint Commission Journal on Quality Improvement, 23*, 135–147.
Song, Y., Skinner, J., Bynum, J., Sutherland, J., Wennberg, J. E., & Fisher, E. S. (2010). Regional variations in diagnostic practices. *New England Journal of Medicine, 363*, 45–53.
Spiegelhalter, D., Sherlaw-Johnson, C., Bardsley, M., Blunt, I., Wood, C., & Grigg, O. (2012). Statistical methods for healthcare regulation: Rating, screening and surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 175*, 1–47. doi:10.1111/j.1467-985X.2011.01010.x
van den Bosch, W. F., Kelder, J. C., & Wagner, C. (2011). Predicting hospital mortality among frequently readmitted patients: HSMR biased by readmission. *BMC Health Services Research, 14*, 57. doi:10.1186/1472-6963-11-57.
van Gestel, Y. R., Lemmens, V. E., Lingsma, H. F., de Hingh, I. H., Rutten, H. J., & Coebergh, J. W. (2012). The hospital standardized mortality ratio fallacy: A narrative review. *Medical Care, 50*, 662–667.
Welch, H. G., Sharp, S. M., Gottlieb, D. J., Skinner, J. S., & Wennberg, J. E. (2011). Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *Journal of the American Medical Association, 305*, 1113–1118.
Wennberg, J. E., Staiger, D. O., Sharp, S. M., Gottlieb, D. J., Bevan, G., McPherson, K., et al. (2013). Observational intensity bias associated with illness adjustment: Cross-sectional analysis of insurance claims. *British Medical Journal, 346*, 1–11. f549. doi:10.1136/bmj.f549

Authors' Biographies

Simon Mackenzie is Medical Director for Quality Improvement NHS Lothian and Clinical Lead for Business Intelligence at Healthcare Improvement Scotland. He was a Health Foundation Quality Improvement Fellow at IHI from 2012 to 2013. He has been a Consultant in Critical Care since 1992, and was Medical Director for University Hospitals in Edinburgh from 2007 to 2013.

Don Goldmann is Chief Medical and Scientific Officer at the Institute for Healthcare Improvement, Professor of Immunology and Infectious Diseases and Epidemiology, Harvard School of Public Health, and Clinical Professor of Pediatrics, Harvard Medical School. He founded and codirects a pediatric health services research training program (T32) at Harvard, which has trained more than 100 fellows over a 20-year period.

Rocco Perla is Assistant Professor of Biostatistics and Health Services Research at the University of Massachusetts Medical School. He is founding Director of the Regional Innovation Network at the Centers for Medicare & Medicaid Services where he previously served as Director of the Learning and Diffusion Group in the Center for Medicare & Medicaid Innovation.

Gareth Parry is Senior Scientist at the Institute for Healthcare Improvement and Clinical Assistant Professor at Harvard Medical School. He is responsible for the rapid-cycle evaluation system and Improvement Capability at IHI.

For more information on this article, contact Gareth Parry at gparry@ihi.org